สุรศักดิ์ สงวนพงษ์

จตุรัสเทคโนโลยี ตอน ดิสค์อะเรย์ เทคนิค

เป็นเทคนิคที่ใช้ในการสร้างดิสค์อะเรย์ โดยจำแนกความแตกต่างได้ถึง 4 รูปแบบ

การพัฒนาความจุของฮาร์ดดิสค์ในปัจจุบันยังมีข้อจำกัดอยู่มาก ฮาร์ดดิสค์ความจุสูงที่เราพบเห็นอยู่โดยทั่วไปจะมีการเชื่อมต่อแบบ SCSI และมักจุข้อมูลได้ราว 1-3 กิกะไบต์ ส่วนฮาร์ดดิสค์ที่มีความจุสูงกว่านี้ก็จะพบเห็นได้น้อยและเริ่มหาได้ยากขึ้นเป็นลำดับ ระบบไมโครคอมพิวเตอร์ที่ต้องเก็บข้อมูลปริมาณมากจึงนิยมที่จะใช้ฮาร์ดดิสค์หลาย ๆ ตัวต่อเชื่อมเข้าด้วยกันเป็น ดิสค์อะเรย์ (Disk Array) เพื่อให้ได้ความจุตามที่ต้องการแทนการใช้ฮาร์ดดิสค์ความจุสูงเพียงตัวเดียว

หากพิจารณาถึงความหมายโดยทั่วไปแล้ว ดิสค์อะเรย์ก็คือกลุ่มของดิสค์หลาย ๆ ตัวที่ต่อเชื่อมอยู่ด้วยกัน ทำให้ได้ดิสค์ที่มีความจุและสมรรถนะสูงกว่าการใช้ดิสค์เดี่ยว ๆ เทคนิคที่ใช้สร้างดิสค์อะเรย์ในปัจจุบันอาจจำแนกตามความแตกต่างได้ 4 รูปแบบ คือ

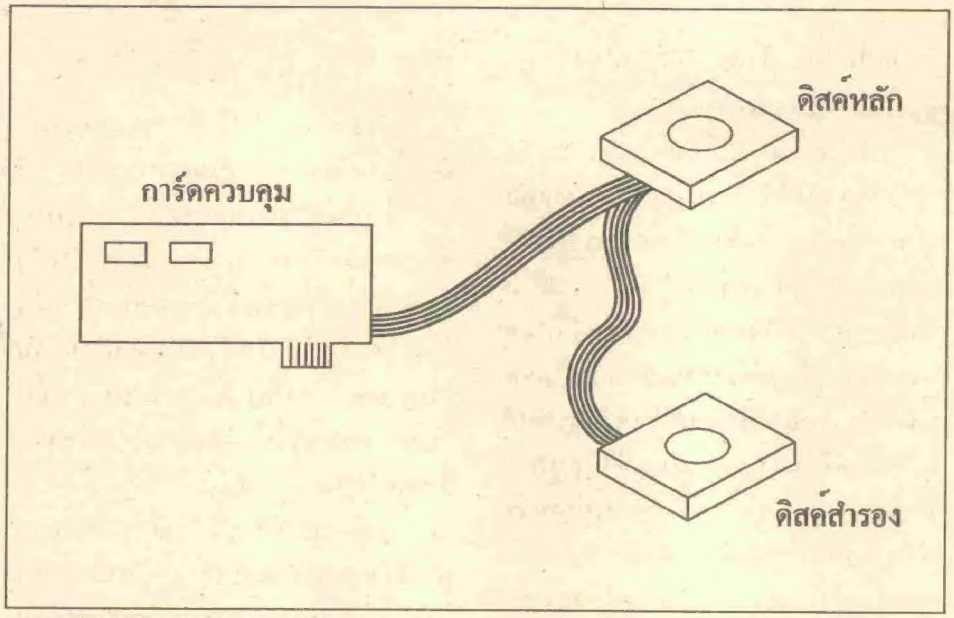
- ดิสค์มิเรอร์ (Disk Mirroring)
- ดิสค์ดูเพล็กซ์ (Disk Duplexing)
- มัลติเปิ้ลไดรฟ์ (Multiple Drives)
- เรด (RAID)

ดิสค์มิเรอร์

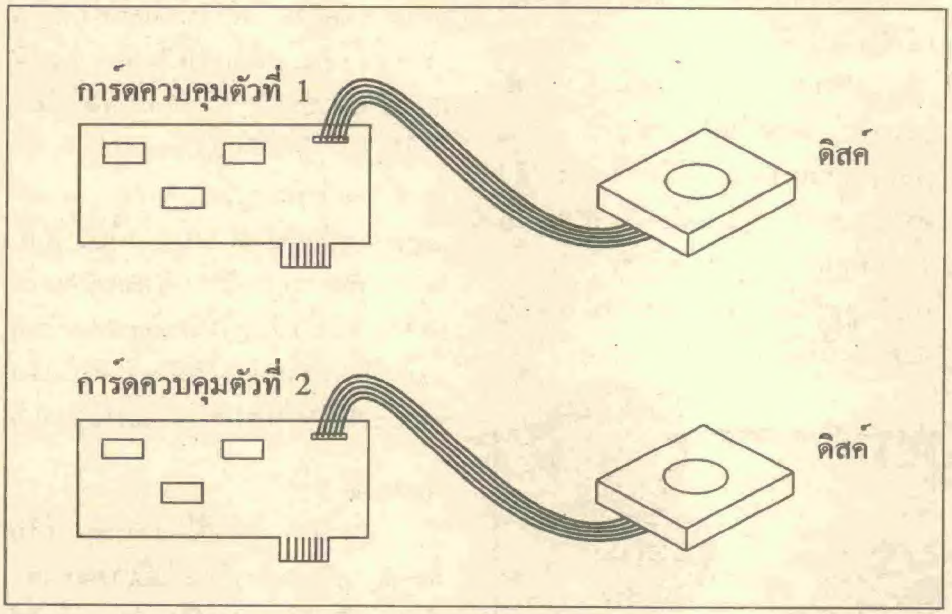
ดิสค์มิเรอร์เป็นเทคนิคการทำดิสค์อะเรย์แบบเบื้องต้นที่สุด หลักการของดิสค์มิเรอร์ก็เพียงแต่ต่อดิสค์สองตัวโดยใช้ตัวควบคุมดิสค์เพียงตัวเดียว ดังตัวอย่างในรูปที่ 1 ข้อมูลจะถูกเขียนลงดิสค์เหมือนกัน ทำให้ดิสค์ทั้งคู่มีข้อมูลเหมือนกัน หรือเปรียบเสมือนเป็นกระจกเงาสะท้อนข้อมูลซึ่งกันและกัน

ประโยชน์ที่ได้จากเทคนิคของดิสค์มิเรอร์ซึ่งมองเห็นได้อย่างชัดเจนได้แก่ ดิสค์ตัวหนึ่งจะทำหน้าที่เป็นตัวสำรองข้อมูล หากเกิดความผิดปกติที่ดิสค์ตัวใดตัวหนึ่ง เราก็คาดหวังได้ว่าจะสามารถกู้ข้อมูลกลับมาได้ทั้งหมดจากดิสค์อีกตัวหนึ่ง เทคนิคของดิสค์มิเรอร์มีทั้งที่เป็นฟังก์ชันภายในของระบบบางระบบเองอย่างเช่นเน็ตแวร์ หรือจำเป็นต้องใช้ซอฟต์แวร์ไดรเวอร์พิเศษเพิ่มเติมเพื่อนำมาใช้กับระบบปฏิบัติการอย่างเช่นเอ็มเอสดอสหรือยูนิกซ์ เพื่อให้สามารถทำดิสค์มิเรอร์ได้

หากมองโดยเฉพาะเพียงเปลือกนอกแล้ว การใช้ดิสค์มิเรอร์ไม่ได้มีความ



รูปที่ 1 ดิสค์มิเรอร์



รูปที่ 2 ดิสค์ดูเพล็กซ์

แตกต่างไปจากการสำรองข้อมูลด้วยวิธีการใช้เทป ยิ่งไปกว่านั้นดิสค์มิเรอร์ยังมีราคาแพงกว่าการใช้เทปอีกด้วย เนื่องจากราคาดิสค์ย่อมแพงกว่าเทป แต่การใช้ดิสค์มิเรอร์เป็นระบบสำรองข้อมูลก็มีข้อดีเหนือกว่าการใช้เทปในแง่ของความเร็วในการสำรองข้อมูล และโดยเฉพาะอย่างยิ่งการเรียกคืนข้อมูลจะทำได้อย่างรวดเร็วกว่าการใช้เทป ซึ่งมักจะกินเวลานับชั่วโมงในการเรียกคืนข้อมูลเพียงไม่กี่ร้อยเมกะไบต์ ดิสค์มิเรอร์จึงเป็นทางออกประการหนึ่งในการเก็บสำรองข้อมูลสำหรับระบบธุรกิจ หรืองานที่มีความสำคัญซึ่งจำเป็นต้องใช้ข้อมูลอยู่เกือบตลอดเวลา

เมื่อมีความผิดปกติเกิดขึ้นที่ดิสค์ตัวใดตัวหนึ่ง ดิสค์มิเรอร์ช่วยให้เราสามารถสลับไปใช้ดิสค์อีกตัวหนึ่งได้ในระยะเวลาอันสั้น โดยการปิดระบบแล้วนำดิสค์ที่มีปัญหาออกมา จากนั้นจึงเปิดระบบคอมพิวเตอร์ใหม่ แม้กระนั้นก็ตามผู้ดูแลระบบก็ยังจำเป็นที่จะต้องสำรองข้อมูลในรูปแบบอื่นอย่างเช่นเทปเข้ามาร่วมด้วย เพราะยังมีกรณีอื่นอีกหลายกรณีที่ก่อให้เกิดความผิดปกติของข้อมูลจนไม่สามารถกู้คืนได้ เป็นต้นว่าระบบไฟฟ้าเกิดผิดปกติหรือเกิดไฟกระชากจนทำให้ดิสค์หลักเสียหายจนใช้การไม่ได้ และดิสค์สำรองก็มักจะเสียหายตามไปด้วย หรือการ์ดควบคุมดิสค์อาจทำงานผิดปกติ ทำให้เขียนข้อมูลทับข้อมูลเดิม กรณีนี้ก็จะเกิดกับดิสค์ทั้งสองตัวเช่นกัน นอกไปจากนี้ความผิดพลาดจากการใช้งาน อย่างเช่น การลบแฟ้มโดยไม่ได้ตั้งใจ การเก็บหรือเปลี่ยนแปลงแฟ้มข้อมูลผิดพลาด ก็เป็นสาเหตุของการสูญหายของข้อมูลในดิสค์ทั้งคู่ด้วย เนื่องจากว่าดิสค์มิเรอร์ใช้หลักการเขียนข้อมูลลงดิสค์ทั้งคู่เหมือนกัน

ดิสค์ดูเพล็กซ์

ดิสค์ดูเพล็กซ์มีหลักการพื้นฐานที่คล้ายคลึงกับดิสค์มิเรอร์อยู่มาก กล่าวคือ มีดิสค์สองตัวทำงานขนานกันไป ข้อมูลที่ถูกเขียนลงบนดิสค์ตัวหนึ่งก็จะถูกเขียนลงบนดิสค์อีกตัวหนึ่งด้วย ส่วนที่แตกต่างออกไปก็คือ ดิสค์ดูเพล็กซ์จะใช้การ์ดควบคุมสองตัว แต่ละการ์ดจะควบคุมดิสค์แต่ละตัวโดยอิสระแยกออกจากกันไป ดังแสดงในรูปที่ 2

เราจะพบว่าในเน็ตแวร์หรือยูนิกซ์จะมีฟังก์ชันที่สนับสนุนการใช้งานดิสค์ดูเพล็กซ์ ในขณะที่ระบบปฏิบัติการอย่างเอ็มเอสดอสไม่ได้ออกแบบมาให้ใช้งานร่วมกับเทคนิคแบบดิสค์ดูเพล็กซ์นี้ได้

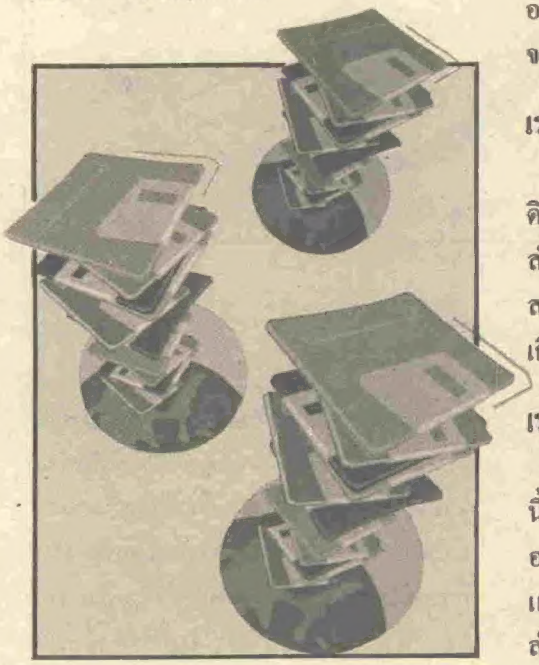
ดิสค์ดูเพล็กซ์จะมีสมรรถนะการทำงานสูงกว่าดิสค์มิเรอร์ เนื่องจากว่าแต่ละการ์ดจะเป็นเสมือนช่องทางสองช่องให้ข้อมูลวิ่งผ่าน แทนที่จะมีลักษณะเป็นคอขวดแบบช่องเดียวเหมือนอย่างในดิสค์มิเรอร์ ข้อได้เปรียบอีกประการหนึ่งของดิสค์ดูเพล็กซ์ที่มีเหนือดิสค์มิเรอร์ คือการแยกการ์ดออกจากกัน ทำให้ความผิดปกติของการ์ดหนึ่งมักไม่ส่งผลกระทบไปยังการ์ดอีกตัวหนึ่ง ดิสค์ดูเพล็กซ์จึงให้ความมั่นใจในการเก็บข้อมูลได้มากกว่าและมีสมรรถนะดีกว่าดิสค์มิเรอร์ แต่ราคาของระบบย่อมสูงกว่าดิสค์มิเรอร์ เนื่องจากต้องเพิ่มการ์ดควบคุมเข้ามาอีกหนึ่งการ์ด

มัลติเปิ้ลไดรฟ์

ในระบบฐานข้อมูลขนาดใหญ่อาจใช้วิธีการจัดเก็บข้อมูลโดยการแบ่งฐานข้อมูลออกเป็นส่วน ๆ แล้วกระจายการจัดเก็บบนดิสค์หลาย ๆ ตัว การจัดเก็บแบบนี้มีหลักการเช่นเดียวกันกับระบบการทำงานพร้อมกันหลายอย่าง (Multitasking) เราจึงเรียกการจัดเก็บด้วยเทคนิคนี้ว่ามัลติเปิ้ลไดรฟ์ ข้อได้เปรียบที่เห็นได้อย่างชัดเจนจากเทคนิควิธีการนี้ก็อย่างเช่น เมื่อคอมพิวเตอร์ต้องการอ่านข้อมูลซึ่งอยู่ในดิสค์ต่างตัวกัน การอ่านข้อมูลจะเกิดขึ้นในดิสค์ทั้งสองตัวได้พร้อมกัน เมื่อเปรียบเทียบกับระบบดิสค์เดี่ยวแล้วจะเห็นได้ว่า คอมพิวเตอร์จะต้องรอให้มีการอ่านข้อมูลส่วนแรกเสร็จเสียก่อน จึงจะให้มีการอ่านข้อมูลส่วนที่สองตามมาได้

ในระบบการเชื่อมต่อแบบ SCSI เราสามารถใช้ตัวควบคุมดิสค์เพียงตัวเดียวสำหรับอ่านเขียนดิสค์หลายตัวได้ แต่ถ้าหากเป็นการเชื่อมต่อแบบ IDE ก็จำเป็นต้องใช้ตัวควบคุมหลายตัวคอยควบคุมดิสค์ เทคนิคแบบมัลติเปิ้ลไดรฟ์มีข้อได้เปรียบเหนือดิสค์มิเรอร์และดิสค์ดูเพล็กซ์ในแง่ของความจุที่ได้ต่อการลงทุน เนื่องจากฮาร์ดดิสค์ทุกตัวใช้สำหรับเก็บข้อมูลด้วยเทคนิคการกระจายข้อมูล ส่วนดิสค์มิเรอร์และดิสค์ดูเพล็กซ์เป็นเพียงเทคนิคของการสำรองข้อมูล ความจุของข้อมูลที่เก็บได้จริงจะเป็นเพียงครึ่งหนึ่งของความจุรวมเท่านั้น

นอกเหนือไปจากเทคนิคของดิสค์อะเรย์แบบมัลติเปิ้ลไดรฟ์นี้แล้ว ยังมีเทคนิคอีกแบบหนึ่งซึ่งโดยผิวเผินแล้วอาจมองดูคล้ายกับเทคนิคแบบมัลติเปิ้ลไดรฟ์อยู่มาก หากแต่มีวิธีการจัดการกับข้อมูลที่แตกต่างกันออกไป เทคนิคดังกล่าวนี้ได้แก่ เรด



เรด

ถึงแม้ว่าการนำดิสค์หลายตัวมาต่อเชื่อมเข้าด้วยกันเป็นแบบอะเรย์ มีข้อดีที่ทำให้ความจุของดิสค์รวมสูงมากขึ้น หากแต่ก็ส่งผลกระทบตามมาให้กับระบบเก็บข้อมูลหลายประการ เช่น โอกาสที่ดิสค์ตัวใดตัวหนึ่งจะเกิดความเสียหายจะมากขึ้น ความร้อนที่เกิดขึ้นในระบบจะสูงขึ้น แหล่งจ่ายไฟต้องทำงานหนักขึ้น เป็นต้น

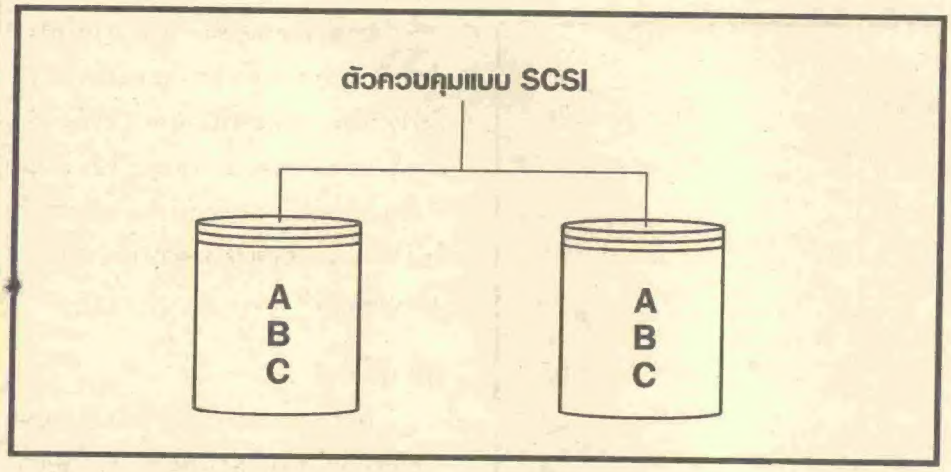
เพื่อป้องกันความเสียหายที่อาจเกิดได้จากผลกระทบต่าง ๆ เหล่านี้ เราสามารถใช้เทคนิคบางประการเพื่อเพิ่มความเชื่อถือได้ให้กับระบบการเก็บข้อมูล เทคนิคซึ่งเป็นที่รู้จักและใช้กันอย่างแพร่หลายได้แก่ เทคนิคของ เดวิด แพทเทอร์สัน, การ์ท กิบสัน และ แรนดี้ แคทซ์ แห่งมหาวิทยาลัยแคลิฟอร์เนียร์ที่เบิร์คลีย์ ซึ่งตีพิมพ์บทความ "The Case for Redundant Arrays of Inexpensive Disks (RAID)" เมื่อปี พ.ศ. 2530 บทความนี้อธิบายถึงเทคนิคแบบต่าง ๆ 5 แบบในการจัดวางระบบดิสค์อะเรย์โดยการแบ่งออกเป็นระดับ เริ่มจากระดับที่ 1 ไปจนกระทั่งถึงระดับที่ 5

เรดระดับ 1

มีหลักการเก็บข้อมูลเหมือนกับดิสค์มิเรอร์ กล่าวคือ ใช้ดิสค์สองตัวสำหรับเก็บข้อมูลที่เหมือนกัน เพื่อให้สามารถกู้ข้อมูลได้เมื่อดิสค์ตัวใดตัวหนึ่งเกิดเสียหาย ดังรูปที่ 3

เรดระดับ 2

วิธีการเก็บข้อมูลในระดับที่สองนี้จะมีการแตกข้อมูลที่ต้องเขียนลงดิสค์ออกเป็นบิต แล้วกระจายบิตนี้ไปยังดิสค์แต่ละตัว การจัดเก็บข้อมูลแบบนี้โดยลำพังแล้วจะเห็นได้ว่ามีความเสี่ยงอยู่



รูปที่ 3 เรดระดับ 1

สูงมาก เนื่องจากหากดิสค์ตัวใดตัวหนึ่งเกิดความเสียหายก็จะไม่มีทางกู้ข้อมูลกลับมาได้เลย เพราะแต่ละส่วนของข้อมูลจะอยู่อย่างกระจายบนดิสค์ต่างตัวกัน การป้องกันปัญหาจึงต้องอาศัยดิสค์เพิ่มเติมเข้ามาเป็นตัวเก็บสำรองข้อมูลสำหรับช่วยในการกู้ข้อมูล หรือเรียกได้ว่าเป็นดิสค์สำหรับตรวจสอบข้อมูล ในดิสค์ตัวตรวจสอบนี้แทนที่จะเก็บข้อมูลต้นฉบับจริงก็จะใช้เทคนิคการเข้ารหัสข้อมูลเพื่อลดขนาดข้อมูลให้เล็กลง รหัสที่ใช้เก็บสำหรับเป็นรหัสแฮมมิ่ง (Hamming code) ซึ่งเป็นรหัสที่มีคุณสมบัติช่วยในการแก้ไขข้อมูลได้ทันทีหากข้อมูลเกิดผิดปกติ ในเรดระดับที่ 2 นี้หากดิสค์ตัวใดตัวหนึ่งผิดปกติ หรือแม้แต่ดิสค์สองตัวเกิดผิดปกติพร้อมกัน ก็จะสามารถกู้ข้อมูลกลับมาได้

เรดระดับที่ 2 นับว่ามีราคาแพง เนื่องจากต้องใช้ดิสค์เพิ่มเติมหลายตัวสำหรับเก็บรหัส ตัวอย่างเช่น เมื่อมีดิสค์ 7 ตัว ดิสค์ 4 ตัวแรกจะใช้เป็นตัวเก็บข้อมูล และ 3 ตัวที่เหลือต้องนำมาใช้เป็นตัวเก็บรหัส ดังรูปที่ 4

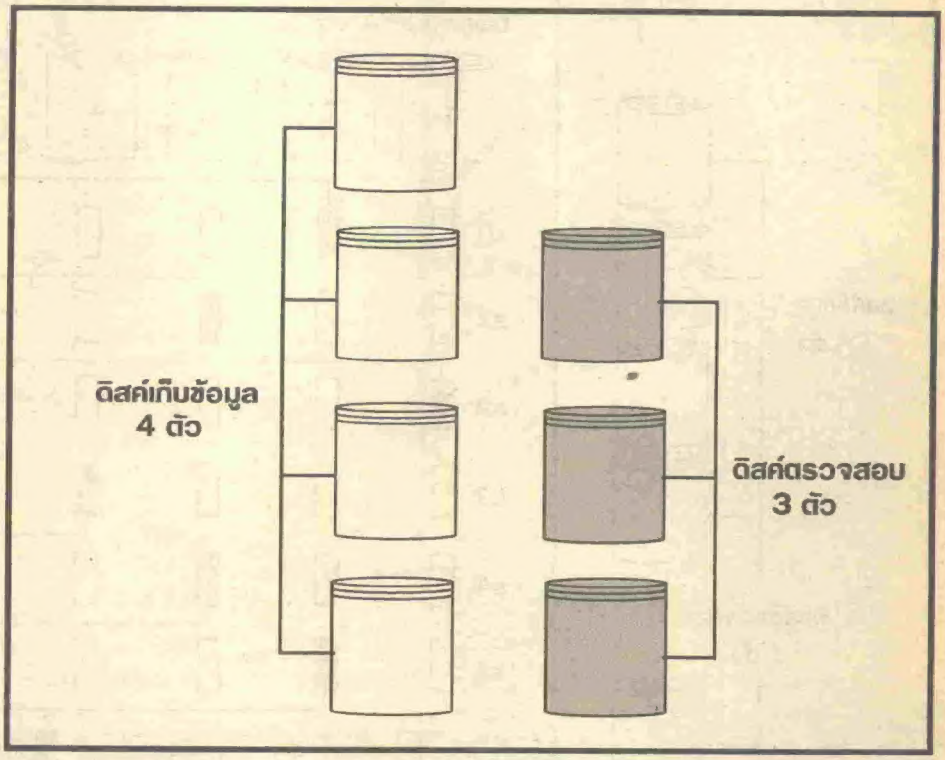
เรดระดับ 3

เพื่อลดจำนวนดิสค์ที่ใช้เก็บรหัสสำหรับกู้ข้อมูลลงไป เราสามารถแตกข้อมูลออกเป็นไบต์แล้วจัดเก็บลงในดิสค์แต่ละตัว เทคนิคนี้จะจัดเป็นดิสค์อะเรย์ระดับ 3 จำนวนดิสค์ที่ใช้เก็บรหัสก็จะลดลงไป สำหรับเรดระดับ 3 ที่มีดิสค์เก็บข้อมูล 4 ตัว จะใช้ดิสค์เพียงตัวเดียวสำหรับจัดเก็บรหัสแฮมมิ่ง ดังรูปที่ 5

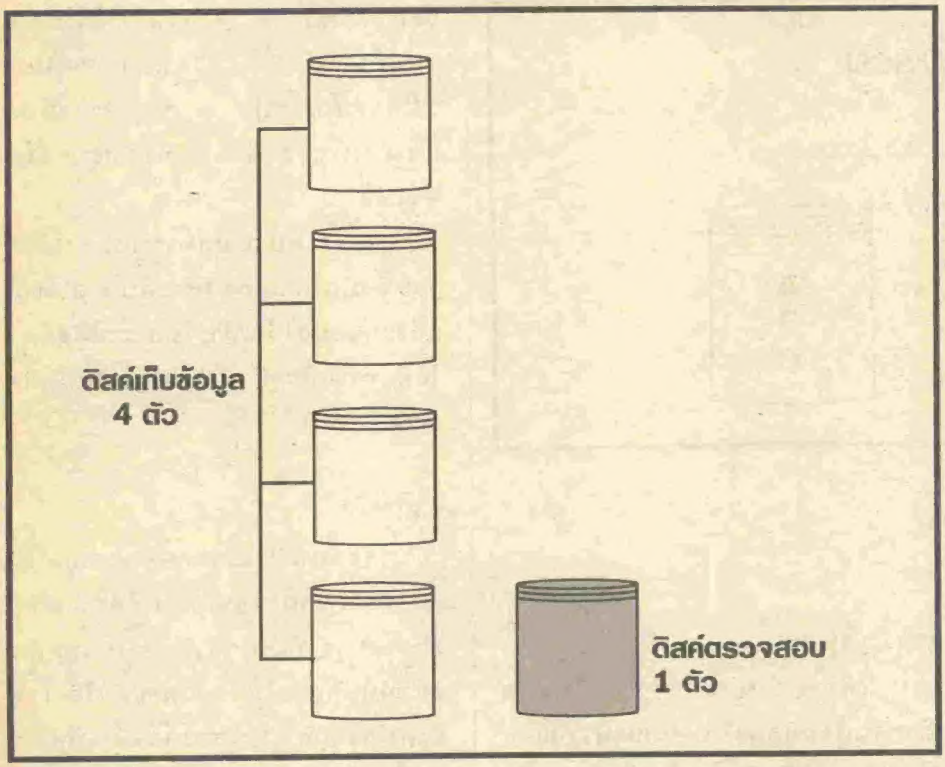
เรดระดับ 3 มีสมรรถนะการถ่ายโอนข้อมูลใกล้เคียงกับระดับที่ 2 หากแต่จะกู้ข้อมูลได้เฉพาะในกรณีที่ดิสค์ตัวใดตัวหนึ่งเพียงตัวเดียวเท่านั้นที่เกิดผิดปกติ

เรดระดับ 4

ระดับที่ 4 จะแตกข้อมูลออกเป็นเซกเตอร์แล้วกระจายลงไปที่ดิสค์แต่ละตัว จึงช่วยให้มีการอ่านเขียนข้อมูลกันอย่างเป็นอิสระในดิสค์แต่ละตัวได้ โดยยังคงใช้ดิสค์เก็บรหัสตรวจสอบเพียง 1 ตัวสำหรับดิสค์ 4 ตัวเช่นเดียวกับระดับที่ 3



รูปที่ 4 เรดระดับ 2



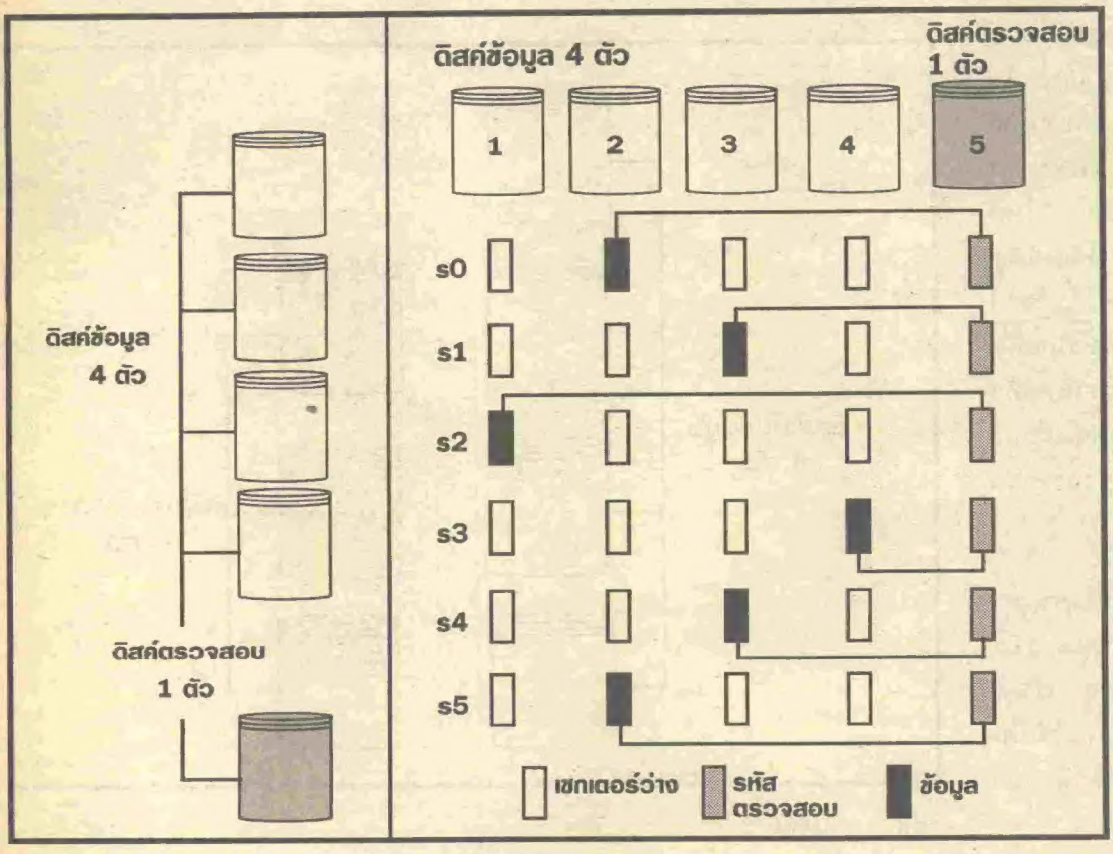
รูปที่ 5 เรดระดับ 3

เราจะสังเกตเห็นได้ว่า หากมีการอ่านเขียนข้อมูลเพียงเซกเตอร์เดียวก็จะมีการติดต่อกับดิสค์ที่เก็บข้อมูลเพียงตัวเดียว และรหัสตรวจสอบจะไปกระจุกตัวอยู่ที่ดิสค์ตรวจสอบเพียงตัวเดียว ทำให้เกิดคอขวดของการอ่านเขียนขึ้นในระบบดังรูปที่ 6

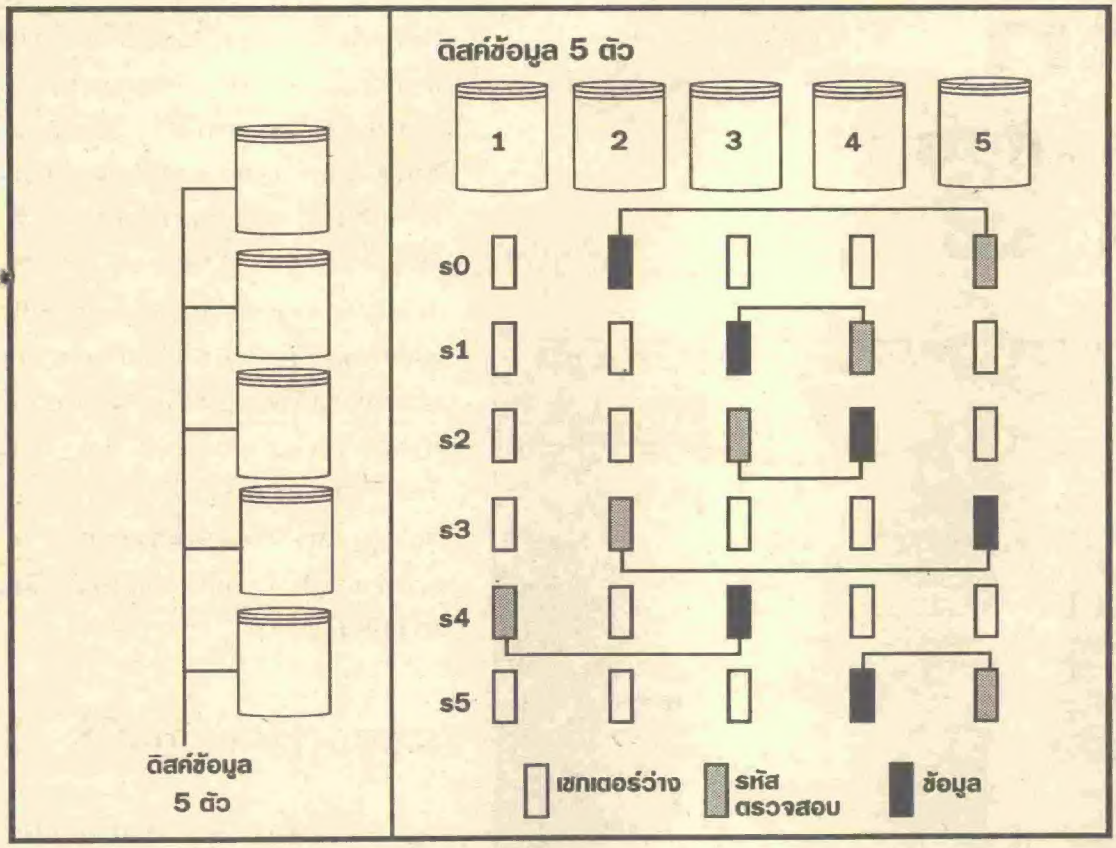
เรดระดับ 5

เทคนิคในระดับที่ 5 ช่วยแก้ปัญหาคอขวดที่เกิดขึ้นในระดับที่ 4 โดยไม่เจาะจงให้ดิสค์ตัวใดตัวหนึ่งเป็นตัวตรวจสอบข้อมูล หากแต่ข้อมูลตรวจสอบจะกระจายตัวไปในดิสค์แต่ละตัวได้ดังรูปที่ 7 และการอ่านเขียนยังคงเป็นแบบเซกเตอร์เช่นเดียวกับระดับที่ 4

เรดระดับ 5 จำเป็นต้องใช้ตัวควบคุมดิสค์แบบพิเศษที่สมรรถนะสูง ซึ่งในปัจจุบันยังมีราคาแพงอยู่มาก



รูปที่ 6
เรดระดับ 4



รูปที่ 7
เรดระดับ 5

ส่งท้าย

จุดสำคัญที่ควรพิจารณาในการนำเอาดิสค์อะเรย์แบบเรดมาใช้คือ สมรรถ-ที่จะต้องเข้าคู่กันกับระบบปฏิบัติการที่ใช้งานอยู่ เทคนิคของเรดจะให้สมรรถนะต่ำอย่างมากเมื่อใช้กับระบบปฏิบัติการอย่างเอ็มเอสดอส หากแต่จะให้สมรรถนะที่ดีสำหรับระบบปฏิบัติการที่ทำงานได้หลายชิ้นงานอย่างเช่นยูนิกซ์ ยิ่งไปกว่านั้นลักษณะของฮาร์ดแวร์ก็ยังมีความสำคัญอย่างมาก อย่างเช่น สล้อตถ่ายโอนข้อมูลตามแบบ ISA ในระบบพีซีมีความเร็วการถ่ายโอนข้อมูลที่ต่ำจึงไม่เหมาะที่จะนำดิสค์อะเรย์แบบเรดมาใช้งาน

โดยความเป็นจริงแล้วสมรรถนะด้านการถ่ายโอนข้อมูลที่เพิ่มขึ้นด้วยเทคนิคของเรดไม่นับว่าเป็นจุดเด่นมากนัก แต่สิ่งที่สำคัญยิ่งกว่านั้นได้แก่ระบบสำรองข้อมูลอัตโนมัติภายในตัว และผู้ใช้สามารถสร้างระบบเก็บข้อมูลความจุสูงได้จากการนำดิสค์ความจุต่ำหลายตัวมาต่อเชื่อมโยงกัน จนกระทั่งได้ความจุเท่าที่ต้องการ ซึ่งอาจจะไม่สามารถหาได้จากดิสค์เดี่ยว ๆ เพียงตัวเดียว

ดิสค์อะเรย์ที่มีประสิทธิภาพสูงก็ย่อมมีราคาแพงตามขึ้นไปด้วย เฉพาะเรดระดับ 5 เท่านั้นที่กล่าวได้ว่าให้สมรรถนะต่อการลงทุนที่นับว่าเหมาะสม ในขณะที่เรดระดับที่ต่ำกว่าจะดีกว่าเฉพาะในแง่ของสมรรถนะของดิสค์โดยตรง ยิ่งไปกว่านั้นเรดแต่ละระดับจะเหมาะสมกับลักษณะของงานประยุกต์ที่แตกต่างกันออกไป ทั้งนี้เนื่องจากเทคนิควิธีการแตกข้อมูลที่ต่างกัน อย่างเช่น เรดระดับที่ 2 และ 3 จะเหมาะสมกับงานประยุกต์ด้านวิทยาศาสตร์ เนื่องจากมีการอ่านเขียนข้อมูลในระดับบิตและไบต์ แต่จะไม่เหมาะสมกับงานประยุกต์เชิงธุรกิจที่ต้องประมวลผลข้อมูลเป็นรายการ (Transaction) ในกรณีดังกล่าวนี้เรดระดับ 4 หรือ 5 จะเหมาะสมยิ่งกว่า